

El tribalismo digital, entre la furia y la farsa: pinchemos la burbuja de la polarización artificial en internet¹

Pedro Jesús Pérez Zafrilla² 

En los últimos años ha crecido la preocupación en la opinión pública sobre la polarización política. Algunos autores señalan internet y las redes sociales como una causa de esa polarización. En ese contexto, este artículo aborda los diferentes usos que los sujetos hacen de internet y que dan lugar a distintas formas de polarización. El artículo parte de la revisión bibliográfica sobre las bases neuropsicológicas del comportamiento político y la naturaleza de la polarización política. Sobre este fundamento, el artículo presenta las formas en que tiene lugar la polarización en internet, así como los supuestos morales y epistémicos de la polarización política. Frente a estas formas de polarización, el artículo desarrolla el concepto de *polarización artificial*. Este concepto pretende explicar cómo los usos expresivos de la red generan una percepción errónea de polarización entre los usuarios. Ejemplos de polarización artificial son los discursos incendiarios, los linchamientos virtuales o el exhibicionismo moral. Como resultado, el artículo presenta algunas indicaciones para desactivar los procesos de polarización artificial y construir un ambiente más sosegado en la red.

Palabras clave: polarización política; polarización artificial; discursos incendiarios; linchamiento virtual; exhibicionismo moral

Introducción

En los últimos años, diversos estudios realizados tanto en EE. UU. (Pew Research Center, 2017) como en Europa (Ipsos Mori, 2018) alertan del aumento de la polarización política en la sociedad. Entre las causas de este fenómeno, son varios los autores (Stroud, 2011; Haidt y Rose-Stockwell, 2019) que señalan a internet y al uso de redes sociales. Las comunicaciones entre afines y los algoritmos que facilitan la búsqueda de información que confirma a los sujetos en sus posiciones estarían provocando una mayor polarización de

¹ Este estudio se inserta en el Proyecto de Investigación Científica y Desarrollo "Ética cordial y Democracia ante los retos de la Inteligencia Artificial" PID2019-109078RB-C22, financiado por MCIN/ AEI /10.13039/501100011033.

² Universidad de Valencia, Departamento de Filosofía, Facultad de Filosofía y Ciencias de la Educación, Valencia, España. E-mail: <p.jesus.perez@uv.es>.

nuestras sociedades, así como la corrosión del debate público y de la convivencia democrática.

Ahora bien, determinar la incidencia de internet en la polarización política no puede hacerse sin atender a los diferentes usos que las personas hacen de la red. Ello es así porque, como trataré de mostrar en este artículo, hay usos de internet que producen una polarización política real, es decir, un reforzamiento de las convicciones que provoca un distanciamiento de los sujetos en sus posiciones políticas. En cambio, existen otros usos que generan lo que he llamado una *polarización artificial*. Con este concepto hago referencia a los usos expresivos de la comunicación mediante los que los usuarios buscan mejorar su estatus en la red recurriendo para ello a una escalada de comentarios inflamantes (Pérez Zafrilla, 2021). Ejemplos de polarización artificial son la exposición a cierto contenido altamente ideologizado, la participación en linchamientos virtuales o el exhibicionismo moral. Estos usos de la red crean en la audiencia una percepción ficticia de polarización que es necesario conocer para así desenmascararla y combatir sus efectos sobre los usuarios.

En este artículo analizo esta distinción entre los usos de internet que producen formas de polarización política frente a los usos de internet que dan lugar a una polarización artificial. En una primera sección titulada "Definir la polarización" abordo el fenómeno de la polarización política desde las tesis de la psicología evolucionista sobre nuestra naturaleza tribal. Además, hago unas clarificaciones introductorias sobre el uso del concepto de polarización política en los ámbitos de la ciencia política y la filosofía. Seguidamente, en la sección "Formas de polarización" expongo los modos en que se produce la polarización política, tanto en el mundo *offline* como *online*. Esto nos permitirá conocer los supuestos morales y epistémicos que dan sentido a los procesos de polarización política en internet. Sobre esta base, en la tercera sección titulada "Polarización artificial" desarrollo la idea de polarización artificial, introducida en un trabajo anterior (Pérez Zafrilla, 2021). En esta sección profundizo en la lógica de la red que da lugar a la polarización artificial. Después presento nuevos rasgos de los principales ejemplos de polarización artificial en internet: los discursos incendiarios, los linchamientos virtuales y el exhibicionismo moral. Finalmente, en la sección "Cómo frenar la polarización artificial" presento unas recomendaciones para afrontar los contextos de polarización artificial con el objetivo de rebajar la atmósfera de crispación en la red.

Definir la polarización

Un estudio adecuado de la polarización política en internet debe partir de la naturaleza tribal del ser humano. Las teorías desarrolladas en el campo de la psicología evolucionista comparten la tesis de que el proceso de hominización se desarrolló en el marco de convivencia en grupos pequeños. Este hecho hizo que el cerebro humano se

configurara adoptando unos códigos de conducta que permitieran la supervivencia de los sujetos en esos grupos pequeños. Dichos códigos fomentaban la cooperación con los sujetos cercanos, como forma de evitar el aislamiento o el castigo del grupo. Por ese motivo, desde la psicología evolucionista se reduce el fenómeno moral al comportamiento dirigido a la cohesión social. Los teóricos, sin embargo, difieren en el enfoque que adoptan en su explicación. Unos autores, como Trivers (1971) o Alexander (1987), son partidarios de la selección a nivel individual: la evolución pertrechó a los individuos con las herramientas necesarias, especialmente las emociones morales de vergüenza o indignación, para tener una disposición a la cooperación social y el respeto a las normas del grupo. Ello permitirá a los sujetos mantener su reputación como buenos cooperadores y garantizar así su supervivencia en el grupo. En cambio, otros autores, como Wilson (1995) o Haidt (2019), adoptarán una posición distinta, denominada como teoría de la evolución de grupos: para estos autores los grupos promueven en los sujetos actitudes favorables a la adhesión tribal, como la lealtad y el sacrificio. De esta forma, los grupos que tienen miembros más fieles se imponen evolutivamente sobre los grupos formados por sujetos más egoístas y menos comprometidos con los valores del grupo.

Así pues, la psicología evolucionista, desde diferentes marcos teóricos, nos muestra cómo la evolución, mediante la convivencia durante milenios en grupos pequeños, configurará nuestra naturaleza tribal. Como muestran también los experimentos realizados en el campo de la neurociencia (Greene, 2013), el cerebro del ser humano está programado para favorecer la cooperación con los cercanos, y, por ende, para la formación de grupos y la adhesión a los mismos como forma de garantizar su supervivencia individual.

Es precisamente esta naturaleza tribal sobre la que se asienta el fenómeno de la polarización política. Pero para comprender adecuadamente la polarización, es preciso realizar una clarificación conceptual. En la actualidad las reflexiones sobre la polarización política en el ámbito académico entremezclan dos definiciones que, aunque relacionadas entre sí, es necesario deslindar. Concretamente, en las ciencias sociales y humanidades el término "polarización" adquiere dos significados distintos, uno en ciencia política y otro distinto cultivado en la filosofía.

En ciencia política la polarización política no se entiende como un extremismo en las posiciones, sino como un alineamiento de los ciudadanos en identidades políticas contrapuestas. Es decir, "polarización política" significa división de la sociedad en dos grupos contrapuestos a través de la adhesión acrítica de los ciudadanos a las posiciones de su partido preferido frente a las posiciones defendidas por los adversarios. Más concretamente, en ciencia política podemos definir la polarización política como el "proceso por el que las diversas perspectivas surgidas en la sociedad quedan realineadas en una única dimensión: la pugna entre dos grupos, 'nosotros' frente a 'ellos'" (McCoy, Rahman y Somer, 2018, p. 18). Con ello, la política, en lugar de definirse, en la línea aristotélica, como el ámbito de diálogo para llegar a acuerdos entre ciudadanos, queda configurada

como una lucha agonística entre dos grupos (partidos, en los sistemas bipartidistas, o bloques compuestos de partidos en los sistemas multipartidistas) que encarnarán el bien y el mal absolutos.

Esta definición de polarización política como división de la sociedad en grupos contrapuestos a los que los sujetos se adhieren permite aunar las dos formas en que se estudia este fenómeno en el campo de la ciencia política: la polarización ideológica y la polarización afectiva.

La polarización ideológica se mide a partir del posicionamiento de los sujetos sobre los distintos temas. Una sociedad moderada será aquella en la que una mayoría de ciudadanos muestra una posición independiente sobre los distintos asuntos políticos. Así los sujetos adoptarán una posición conservadora sobre unos temas y una posición más progresista sobre otros, de tal forma que los sujetos no siempre coinciden con la posición de un partido sobre esos temas. En cambio, una sociedad está más polarizada ideológicamente cuando los sujetos alinean su posición sobre los distintos temas con la mantenida por el partido con el que se identifican (Pérez Zafrilla, 2020). Es decir, los sujetos compran de forma acrítica el paquete ideológico de su tribu política al pensar sobre los distintos asuntos, de tal forma que los individuos actuarán guiados por su identidad grupal (Kim, 2016). Como consecuencia, se reforzará la adhesión ideológica de los individuos con los partidos y, con ello, crecerá la división en dos grupos de la sociedad. La polarización política no es, por tanto, un sinónimo de radicalización, sino de división de la sociedad en dos grupos contrapuestos.

Por su parte, la polarización afectiva hace referencia a los sentimientos surgidos hacia el adversario político como consecuencia de la identificación que los sujetos hacen con su partido. En una sociedad moderada los sujetos no se identificarán fuertemente con un partido, al tener posiciones independientes sobre los diversos asuntos. Por tanto, los sentimientos de afecto hacia su propio partido y de rechazo a los adversarios serán bajos. En cambio, donde reina una fuerte identificación partidaria, los sentimientos de lealtad con su partido y de rechazo y animadversión a los adversarios, serán mayores. Esta brecha de sentimientos hacia el propio partido y hacia el adversario indica una mayor polarización afectiva de la sociedad (Iyengar y Westwood, 2015).

La polarización afectiva es el resultado de nuestra naturaleza tribal. La pertenencia a un grupo genera lazos de amistad y cooperación con los miembros del grupo y rechazo hacia los de fuera, al percibirlos como una amenaza para la supervivencia. Esta es la base del conocido sesgo de intragrupo/extragrupo, según el cual, la mera identificación con un grupo genera sentimientos positivos hacia los miembros del grupo y negativos hacia los oponentes. Así, la gente tiende a pensar que su propio grupo busca el bien común, mientras que los adversarios son malvados y se guían por intereses egoístas. Esto hace que los sujetos que se identifican con un partido tiendan a desarrollar sentimientos negativos hacia sus adversarios (Haidt, 2019). Por tanto, cuando en una sociedad existe un alto grado de

rechazo, animadversión o desconfianza hacia los miembros del partido opuesto, se incrementa la polarización afectiva.

Esta identificación de los sujetos con unos partidos enfrentados, unido al rechazo sentido hacia los adversarios, arroja como consecuencia un menor punto de acuerdo y de entendimiento entre la ciudadanía y, con ello, una mayor división de la sociedad. Ello explica que el centro político (entendido como el grupo de población con menor identificación partidaria) pierda apoyo y en su lugar los individuos ideologizados aumenten su número. El resultado es la polarización de la sociedad al quedarse dividida en dos grupos irreconciliables.

Pues bien, en este contexto de polarización política entendida como fuerte división de la sociedad en dos grupos antagónicos, entra en juego el sentido de polarización desarrollado en las humanidades. En filosofía, el término "polarización" se aplica a un fenómeno producido principalmente en el intercambio discursivo entre sujetos. Cuando los sujetos dialogan, o cuando acceden a contenidos altamente ideológicos, como sucede en internet, en determinados contextos se puede producir una polarización de posiciones. En este caso, polarización tampoco significa una tendencia al extremismo o la radicalidad. Se entiende más bien en el sentido de que los sujetos pasan a mantener una posición más convencida en la línea de lo que pensaban antes del diálogo o del acceso a ese contenido ideologizado (Sunstein, 2003). Polarización es, así, un reforzamiento de las convicciones de los sujetos.

En este artículo el término "polarización" se entenderá en este sentido de reforzamiento de posiciones de los sujetos. Pero es importante tener en cuenta que, como mostraré, el reforzamiento de posiciones de los sujetos toma como base la adopción de la identidad grupal por parte de los individuos, y esta adopción de la identidad grupal es fruto de la adhesión del sujeto al paquete ideológico de su partido. Por tanto, la polarización entendida como reforzamiento de posiciones está ligada a la polarización política como adhesión acrítica a la identidad grupal que da lugar a la división de la sociedad en dos grupos. Este es un punto clave que se pondrá de manifiesto a lo largo del artículo pero que es necesario señalar aquí de entrada.

Una vez presentada la idea de polarización política, la cuestión que debemos abordar es cómo se produce esa polarización de posiciones de los sujetos en el mundo digital y que da como resultado esa división de la sociedad.

Formas de polarización

En esta sección desgranaré las formas en que tiene lugar la polarización política entendida como ese reforzamiento de posiciones. Según la bibliografía existente, la polarización de posiciones se produce de dos formas, aplicables tanto al mundo *online* como al *offline*: Hay una polarización intragrupal descrita principalmente por Sunstein

(2003). Esta se produce cuando los miembros de un grupo refuerzan su posición por sí solos o en diálogo con afines. La otra forma de polarización es la intergrupala, señalada desde la neurociencia (Schulz, 2010). Esta tiene lugar cuando dos grupos enfrentados polarizan sus posiciones al confrontar dialécticamente entre ellos. Presentaré en una primera subsección la polarización intragrupal, ya que sienta las bases para comprender la polarización intergrupala, la cual será expuesta en una subsección posterior. Finalmente, dedicaré una última subsección a sintetizar los elementos definitorios de la polarización política entendida como reforzamiento de posiciones.

Polarización intragrupal

La polarización intragrupal se produce de dos maneras: de forma individual y mediante la relación con sujetos de ideas afines. De forma individual, los sujetos identificados con las posiciones de un partido polarizan sus posiciones cuando se informan sobre un asunto en medios de comunicación acordes con las ideas de su partido. Como se apunta desde la psicología social (Kunda, 1990; Shaw, 1996), la búsqueda de información política se rige por un sesgo de confirmación. Este sesgo, por un lado, lleva a los individuos a informarse a través de medios que confirman sus impresiones sobre los asuntos. Por otro lado, el sesgo de confirmación produce la tendencia en los sujetos a evitar aquellos medios que cuestionan esa perspectiva. Ello es así porque los individuos no se informan sobre política para tener una visión objetiva de la realidad. Más bien los individuos tienden a verse como miembros de un partido y quieren encontrar información que apoye a su equipo y refute las posiciones de los adversarios (Brennan, 2018). Para ello los individuos siguen medios especialmente politizados, cuya línea editorial está alineada con el ideario de su partido. El seguimiento de estos canales reforzará la posición de los sujetos sobre los distintos temas en la línea de lo defendido por el partido³. Así se explica que los individuos más informados sobre política tengan posiciones más alineadas con las de su partido que los sujetos menos informados, que son, por ello, más moderados. Hay estudios que reflejan esa conexión entre información política y polarización: cuanto más informado está un sujeto sobre política, mayor es su polarización (Abramowitz y Saunders, 2008). Es decir, cuanto más informada está una persona, más alineada está con la posición de su partido y más convencida está de la verdad de su posición.

Pero si esta polarización de posiciones se produce con los medios tradicionales, en mayor medida podría suceder *online*, gracias a la personalización del algoritmo en las diferentes plataformas digitales (redes sociales, buscadores o agregadores de noticias). El algoritmo puede definirse como un constructo matemático que arroja un resultado

³ De esta manera se aprecia la conexión entre los dos sentidos de polarización existentes en filosofía y ciencia política. La polarización entendida como reforzamiento de posiciones va unida a una mayor adhesión a la tribu política de los sujetos.

determinado a partir de unas variables prefijadas (Hill, 2015). El algoritmo es la herramienta que permite a las plataformas crear un perfil de cada usuario para arrojar los resultados más acordes a lo que el usuario está buscando en los buscadores o lo que más le interesará ver en su muro de la red social (Cho et al., 2020). Esto, en principio, es positivo. Un buscador de internet será eficiente si acierta en seleccionar aquella información de la red que más se acerca a lo que el usuario está buscando en cada momento, y para ello es necesario que la plataforma tenga un perfil sobre sus intereses.

El problema está en que, como denuncian autores como O'Neil (2017) o Pariser (2017), el modelo matemático que está a la base del algoritmo no está diseñado para generar una sociedad más justa ni para proporcionar una visión objetiva de la realidad. El algoritmo está diseñado para aumentar los ingresos, tanto de las plataformas de internet como de las empresas a las que venden nuestros datos. Esta finalidad lucrativa ayuda a entender los tres filtros principales que estas plataformas emplean para seleccionar el contenido ajustado al perfil de cada usuario (Möller et al., 2018):

- a) Popularidad entre el conjunto de usuarios: las recomendaciones más seleccionadas por los usuarios se vuelven más populares y aparecen primero. En cambio, aquellas recomendaciones menos seleccionadas son postergadas en la lista;
- b) Filtro semántico: la plataforma selecciona un contenido en función de los términos de búsqueda empleados antes por el mismo usuario. Así, el algoritmo elige aquellos resultados que se asemejan al espectro de resultados que el usuario seleccionó anteriormente;
- c) Filtro colaborativo: el algoritmo recomienda aquellas entradas que han sido seleccionadas antes por otros usuarios con gustos similares.

Cada plataforma empleará uno de estos filtros o varios de ellos combinados. Por ejemplo, Facebook y Twitter no presentan a los usuarios las publicaciones de sus contactos por orden cronológico, sino aquellas que pueden ser más interesantes al usuario. Para ello el algoritmo de estas plataformas tiene en cuenta factores como las veces que el sujeto ha interactuado con ese contacto o el número de clics que ese post ha recibido (Agrawal, 2016). La prioridad dada a los contenidos que resulten atractivos para los usuarios se debe a que los algoritmos están diseñados para actuar de acuerdo con las tendencias humanas, como sucede con el sesgo de confirmación: los usuarios tienden a preferir los contenidos que confirman sus posiciones previas, evitando aquellos que les resulten disonantes. Por ello, crear algoritmos que dan como resultado información agradable para los usuarios conllevará que estos estén más pendientes de los contenidos que les proporcionan esas plataformas (Cho et al., 2020). Porque, al fin y a la postre, estas plataformas están diseñadas para mantener la atención constante de los usuarios (creando para ello sistemas

de alertas, por ejemplo), ya que este es el modo más eficaz de incrementar los beneficios económicos de las plataformas.

Ahora bien, esta finalidad lucrativa de las plataformas, alimentada por el suministro de contenido agradable a los usuarios, arrojará consecuencias nefastas para la vida política. La principal es que los sujetos pueden caer atrapados en el filtro burbuja. Como denuncia Pariser (2017), en los medios tradicionales la persona tiene conocimiento del conjunto de cadenas de televisión o periódicos que hay de información política y es consciente de cuáles elige y cuáles rechaza seguir. En cambio, en internet el individuo no es consciente de los contenidos que le oculta el algoritmo, ya que solo tiene acceso a los resultados que le devuelve el buscador. Cuando alguien busca contenidos de información política, el algoritmo personalizado devolverá resultados acordes al perfil que la plataforma se ha hecho de la perspectiva e ideología de ese usuario y eliminará de los resultados las perspectivas diferentes sobre esos mismos temas.

Según Pariser, las consecuencias de este escenario sobre la polarización son evidentes. Como los sujetos desconocen la configuración del algoritmo, al informarse sobre política en internet creen encontrar unos resultados que conforman una visión objetiva de los problemas. Pero en realidad el algoritmo oculta de los sujetos las perspectivas disonantes y arroja resultados agradables, que son los coincidentes con su ideología. Por ese motivo, el filtro burbuja hace que los sujetos encuentren información agradable, que les acaba reforzando sus posiciones previas sobre los asuntos públicos. En consecuencia, este reforzamiento hace que los grupos acaben más alejados en sus posiciones sobre los distintos temas, produciéndose así una polarización de la sociedad.

Ahora bien, lo cierto es que la evidencia sobre si la exposición a contenido ideológico en internet aumenta la polarización no es concluyente. Hay estudios que reflejan un aumento de la polarización tras visualizar contenido ideológico en distintas plataformas como YouTube, Twitter o Facebook (Basil et al., 2018; Cho et al., 2020). Otro interesante estudio (Flaxman, Goel y Rao, 2016), basado en los historiales de búsqueda de 50.000 sujetos que se informan de política en internet, encontró que los individuos se informan en la red fundamentalmente por medios acordes con su ideología, y los sujetos partidistas lo hacen en mayor medida que los moderados. Esto hace que estas personas estén expuestas casi exclusivamente a un único lado del espectro político, reforzándose con ello la polarización al confirmarlas en sus posiciones. Pero el estudio determinó que la polarización se debe más a la consulta directa de medios *online* que al papel de las redes sociales. Contrariamente a lo que se suele pensar, las redes proporcionan una mayor heterogeneidad de contenidos que la búsqueda directa en la web de medios *online*. Afortunadamente, el estudio también refleja que los sujetos se informan principalmente a través de la consulta directa en la web de medios *mainstream*, mientras que los medios

más radicales, habituados a difundir noticias falsas o teorías de la conspiración, no tienen una incidencia apreciable en el consumo de noticias⁴.

Sin embargo, otro estudio (Boxell, Geintzkow y Shapiro, 2017) encuentra una correlación inversa entre el uso de internet para informarse sobre política y el nivel de polarización (ideológica y afectiva) por franjas de edad en EE. UU. El estudio concluye que entre 1996 y 2016 creció la polarización en todas las franjas de edad, pero sobre todo entre los mayores de 65 años. De este modo, dado que son los jóvenes los que más se informan de política por internet, la polarización ha crecido sobre todo entre aquella población con menor consumo de información política *online*. En consecuencia, no parece fácil sostener que el aumento de la polarización en EE. UU. sea debido al uso de internet, ya que la correlación entre este uso y la polarización es claramente inversa. En este sentido, para estos autores la polarización vendría más de la mano de los medios de comunicación de masas más consumidos por la gente de mayor edad.

Junto a la visualización de contenido ideológico, el otro uso de internet que genera una polarización intragrupal es el diálogo entre sujetos que comparten unas mismas ideas. Aquí se produce lo que Sunstein denomina polarización de grupo. Este es el fenómeno por el cual cuando los miembros de un grupo con ideas afines hablan entre sí, el diálogo lleva a los sujetos hacia una posición más contundente en la línea de lo que el grupo pensaba en el inicio del debate. Es decir, tras el diálogo los sujetos pasan a mantener una posición más convencida sobre ese tema. La polarización grupal se produce tanto cuando las personas conversan cara a cara como en las conversaciones *online*. Aunque internet facilita que individuos con ideas diversas entren en contacto entre sí, las personas suelen emplear internet para contactar y dialogar con gente que comparte su misma ideología o una misma posición sobre ciertos asuntos o para buscar información que refuerce su punto de vista sobre los mismos temas. Esto hace que los sujetos acaben cayendo en internet en cámaras de eco en las que los sujetos refuerzan (y, en consecuencia, polarizan) sus posiciones (Sunstein, 2002, 2003).

Para Sunstein (2003), las causas de la polarización grupal son tres, y se dan tanto *online* como *offline*. La primera tiene que ver con el conjunto de argumentos disponibles: cuando dialogan entre sí personas que comparten una posición sobre un tema (por ejemplo, el apoyo a la creación de un impuesto a las grandes fortunas), los argumentos que haya a favor de esa medida serán muchos más que los que se puedan exhibir en su contra. Esto hará que los sujetos que comparten el apoyo a ese impuesto encuentren nuevos argumentos para defender su posición, lo que les permitirá apoyarla con más fuerza.

⁴ Este último dato es importante, ya que nos permite concluir que los sujetos saben discernir la credibilidad de los medios que consultan para informarse sobre política y que la consulta de medios radicales es muy minoritaria en la sociedad.

La segunda es relativa a la influencia social: la psicología social ha evidenciado que las personas buscan mantener una buena imagen ante los demás miembros del grupo, así como tener una percepción positiva de sí mismos. Esto hace que las personas ajusten su posición sobre un tema a aquella corriente que perciben como mayoritaria dentro del grupo, con el fin de mantener su reputación dentro del mismo (Sunstein, 2003). Si, por ejemplo, los sujetos perciben que la mayoría del grupo está a favor del impuesto a los ricos, aquellos que lo apoyan más firmemente se sentirán más motivados para presentar sus posiciones. En cambio, aquellos pertenecientes a la minoría que no creen del todo que sea tan buena medida, tenderán a reservarse su opinión para no perjudicar su posición en el grupo. Prefieren quedarse callados y parecer que también apoyan la medida antes que expresar sus dudas sobre ella y causar una mala impresión entre la mayoría de sus compañeros.

La tercera causa hace referencia a la relación entre confianza, extremismo y corroboración: cuanto mayor sea el apoyo que una idea tiene dentro de un grupo, mayor será la confianza con que los sujetos apoyan esa idea. Además, dentro del grupo los sujetos expondrán argumentos en apoyo de sus posiciones que tal vez no presentarían en contextos más heterogéneos o donde esa idea no predomine. Por eso los sujetos acaban reafirmando su posición sobre el asunto, al verla corroborada por más y nuevos argumentos (Sunstein, 2002).

Hasta aquí he abordado cómo se produce la polarización intragrupal. En la siguiente subsección paso a analizar cómo tiene lugar la polarización intergrupala.

Polarización intergrupala

La polarización intergrupala *online* y *offline* tiene lugar cuando los sujetos interaccionan con otros individuos con ideas distintas. Si bien hay autores que argumentan que el diálogo cara a cara entre personas con puntos de vista distintos tiende a conducir a la moderación de posiciones (Mutz, 2006), otros autores defienden que el diálogo entre diferentes individuos puede generar formas de polarización de posiciones. Esto sucede cuando se abordan temas especialmente polémicos, como la política penitenciaria o el aborto. Al dialogar sobre estos temas, los sujetos, lejos de acercar posiciones, acaban manteniendo una opinión más convencida y firme sobre los mismos en la línea de lo que pensaban antes del diálogo y adoptando una visión negativa del otro interlocutor (Schulz, 2010).

Por ello, es un hecho que los procesos de polarización intergrupales existen. Para algunos autores de la neuroética la polarización representa una patología de la deliberación. Así Mercier y Sperber (2011) defienden que cuando el diálogo se produce en unas condiciones adecuadas la polarización no tiene lugar. Esas condiciones adecuadas son tres: La primera es la existencia de puntos de vista diferentes en el diálogo (lo que

distingue propiamente la deliberación del diálogo entre afines señalado por Sunstein (2003)). La segunda es que se produzca una evaluación de los argumentos del interlocutor que lleve a discernir la información valiosa que este pueda aportar. Por último, la tercera es que los sujetos estén orientados a la búsqueda de la verdad y no a la defensa recalcitrante de una posición sin escuchar los argumentos del interlocutor. Dadas estas condiciones, el diálogo puede conducir a los sujetos al acuerdo racional y evitar que se produzca la polarización de posiciones.

Sin embargo, a mi parecer, la polarización intergrupal, lejos de representar una patología de la deliberación, es una respuesta natural ante la falta de acuerdo en el diálogo, y viene causada por las diferencias cognitivas que separan a los sujetos (Pérez Zafrilla, 2017). Es decir, los sujetos cuando deliberan lo hacen sobre unos presupuestos morales y epistémicos básicos (Cortina, 2007): el reconocimiento del otro como interlocutor válido, la expectativa recíproca de convencimiento racional, la simetría de las partes o la expectativa recíproca de estar guiado por un criterio de justicia y no por intereses egoístas. No obstante, sucede que las diferencias ideológicas que separan a las personas hacen que deliberen desde marcos cognitivos inconmensurables. Así, cuando se abordan asuntos que los sujetos consideran fundamentales para su identidad grupal, los mismos consideran que la posición contraria a la suya es intolerable, al amenazar lo que como miembros de un grupo consideran un valor irrenunciable (Haidt, 2019)⁵. A ello se une que, en el diálogo, los individuos suponen que el interlocutor asigna a los términos el mismo sentido que ellos les dan (Habermas, 1998). Pero los términos que en una ontología moral tienen un sentido, en otro marco moral pueden tener otro distinto. Así, las razones que son significativas para un sujeto pueden no serlo para otro. Esto hace que las razones que para uno son contundentes, no resulten convincentes para su interlocutor. En consecuencia, cada uno pensará que su interlocutor, al no resultar convencido por razones que el hablante considera evidentes, o es un estúpido que no comprende la evidencia de los argumentos presentados, o es que en realidad no está moralmente motivado, sino que defiende posiciones egoístas, produciéndose así la polarización de posiciones (Schulz, 2010). Es decir, ambos sujetos están motivados moralmente, pero la imposibilidad de convencer al adversario en aspectos que el sujeto considera fundamentales para su identidad grupal desata las emociones reactivas de indignación y recelo hacia el interlocutor, lo que deriva en la polarización de sus posiciones: el sujeto se reafirma en su posición y siente un rechazo hacia su interlocutor.

⁵ Ejemplos de valores irrenunciables que suelen salpicar el debate público son: el respeto a la bandera para un nacionalista o el respeto al medio ambiente para un ecologista. En ambos casos, la crítica a estos temas (defender la quema de banderas como forma de protesta, en el primer caso, o defender la construcción de una carretera por un parque natural, en el segundo) es percibida por cada uno de estos individuos como un ataque a su propia identidad grupal.

Esta polarización intergrupal se puede producir tanto en relaciones cara a cara como en internet. Por ejemplo, esto sucede cuando dos usuarios de Twitter se enzarzan en una discusión sobre un tema que les divide, pero lo hacen respetando los supuestos morales y epistémicos que rigen la deliberación y a los que antes hice alusión. Tras el diálogo cada cual se reafirma en sus tesis y, al no poder convencer a su interlocutor, ambos piensan que su interlocutor es estúpido o malvado por no aceptar las razones ni reconocer valores que el hablante considera evidentes.

En esta sección he presentado las formas en que se puede producir la polarización en internet, y que tienen su reflejo también en el mundo *offline*. Hay una polarización intragrupal, por la que los sujetos refuerzan sus posiciones al visualizar contenido ideologizado o al interactuar con afines. Esto los lleva a adherirse a la posición del grupo y, con ello, actuar guiados por su identidad grupal. Pero también hay una polarización intergrupal, cuando un sujeto discute con un adversario ideológico sobre un tema que les separa. Entonces, aunque ambos están motivados moralmente, las diferencias ideológicas que les dividen hacen que el acuerdo sea imposible y cada uno se reafirme en las posiciones de su grupo y pase a dudar de la actitud ética de su interlocutor.

Elementos constitutivos de la polarización política

Dicho esto, podemos comprobar que los procesos de polarización, tanto dentro como fuera de internet, tienen en común dos elementos:

En primer lugar, la polarización se produce porque los sujetos adoptan la identidad del grupo, ya sea por la visualización de contenido ideológico, por la discusión con afines o al sentir su identidad grupal dañada al discutir con un adversario. Además, esa identificación con el grupo les hace desarrollar sentimientos de rechazo hacia el adversario, como refleja la polarización afectiva a la que antes hice alusión.

El segundo elemento es que, a pesar de la adopción de la identidad grupal, los sujetos hacen uso de la razón comunicativa. Este es el punto clave. Los sujetos que reciben una información ideologizada por medios que consideran creíbles presuponen que ese contenido es veraz, sincero y orientado por un criterio de justicia. Del mismo modo, al dialogar, con afines o con un adversario, los sujetos respetan los supuestos habermasianos de verdad, veracidad y justicia que deben regir la comunicación (Habermas, 1987), así como el deber rawlsiano de civilidad (Rawls, 1996). Es decir, los ciudadanos dan argumentos que creen sinceramente, están guiados por un criterio de justicia, escuchan a sus interlocutores y atienden sus argumentos. Por eso justamente polarizan sus posiciones: porque asumen que los otros también participan en el diálogo de forma sincera. Así, la polarización intragrupal se produce porque escuchando a los otros afines los sujetos aprenden nuevos argumentos con los que reforzar su punto de vista. Del mismo modo, la polarización intergrupal tiene lugar cuando los sujetos advierten que el interlocutor no se

deja convencer por argumentos que el hablante considera evidentes. Entonces el hablante comienza a dudar de la sinceridad del adversario y genera una actitud negativa hacia él.

Pues bien, la tesis que deseo mantener a continuación es que esta combinación de identidad grupal con el uso de la razón comunicativa es precisamente el criterio que permite distinguir los usos de internet que dan lugar a procesos de polarización de posiciones de otros usos de internet en los que se produce lo que he propuesto denominar *polarización artificial*: un proceso de competencia de los sujetos por acaparar atención y mejorar su estatus en la red mediante una escalada de expresiones emotivas de indignación, que da como resultado un clima de aparente polarización o de indignación (Pérez Zafrilla, 2021). Concretamente, cuando esta acumulación de expresiones emotivas se dirige contra los adversarios ideológicos, se genera en la audiencia una imagen deformada del adversario alejada de cómo es éste en realidad, así como una fuerte sensación de rechazo hacia él. Ambos elementos (el rechazo de la audiencia hacia el adversario y el estereotipo deformante de éste creado por ese uso expresivo de la red) son los causantes de un fenómeno bien conocido en la ciencia política: la polarización percibida, es decir, la sensación ilusoria que tienen las personas de que existe una gran distancia entre ellas (que se perciben como objetivas y justas) y sus adversarios (a los que ven erróneamente como radicales y malvados), de tal forma que piensan que la polarización política es mayor de la que existe en realidad (Hwang, Kim y Huh, 2014). Esta es una percepción errónea de la polarización porque se basa en una caricatura del adversario construida a partir de esa escalada de discursos expresivos en la red. Estos son además unos discursos cuyos autores no pretendían polarizar la sociedad, sino lograr notoriedad mediante la denigración gratuita del adversario o simplemente descargar un sentimiento de indignación contagiado por el uso de la red.

Esta polarización artificial surge cuando la identificación grupal de los sujetos desborda el uso de la razón comunicativa y en su lugar el uso de la comunicación es puramente expresivo, no discursivo. En un trabajo anterior expuse el origen y los supuestos filosóficos y psicológicos de este concepto, así como el proceso que genera la polarización artificial en diferentes usos de la red (Pérez Zafrilla, 2021). En la siguiente sección profundizaré en los rasgos de la polarización artificial en internet.

Polarización artificial

La polarización artificial es fruto de ciertas lógicas existentes en internet. Como explica Han (2014), los medios analógicos (televisión y radio) agrupan a los sujetos diluyéndolos en una masa. En cambio, los medios digitales (internet y redes sociales) aíslan a los individuos presentándolos ante los demás con un perfil propio, que puede ser anónimo. De este modo, en los medios digitales los usuarios se encuentran aislados, convirtiéndose en sujetos ayunos de atención. Esto hace que los usuarios busquen a otros

semejantes con los que cubrir unas necesidades afectivas de reconocimiento y pertenencia. Este es el elemento que explica que en las redes sociales prime la comunicación expresiva sobre el diálogo racional. Por un lado, la gente no emplea la red para dialogar con el otro – su adversario – sino para expresar opiniones, sentimientos y estados de ánimo por un mero desahogo. Además, la comunicación en redes sociales está viciada al quedar sometida al tribunal de la validación social. Toda expresión tiene asignado un estatus en función de los *likes* o *retuits* que recibe cada publicación. Como señala Alter (2018), una publicación sin *likes* no sólo daña al individuo en la privacidad de su casa, sino que se convierte en una condena pública: muestra a ojos de todos que las opiniones de ese sujeto no interesan a nadie. Por ese motivo, el hecho de que la interacción en la red esté sometida a la necesidad de validación de otros usuarios, empuja a los sujetos a buscar constantemente recompensas a sus mensajes, publicando contenidos emocionalmente cargados que atraigan la atención de la audiencia.

Por otro lado, los sujetos se mueven en reacción a los estímulos recibidos, buscando la convergencia emocional con sus semejantes y la reafirmación de su identidad grupal. Se produce así la conectividad por la que los individuos se adhieren de forma instantánea a corrientes de opinión promovidas por la tribu moral con la que se identifican. En consecuencia, las redes, en lugar de procurar una esfera pública deliberativa, generan más bien una atmósfera o estado de ánimo medido por el termómetro del *trending topic*, que los distintos grupos buscan copar frente a sus adversarios. Para ello, los grupos recurrirán a bots, cuentas falsas y difusión de bulos que amplifiquen los mensajes del grupo y opaquen los del adversario (Rubio Núñez, 2018).

En síntesis, el ecosistema de internet potencia tres elementos: el uso expresivo de la comunicación, la búsqueda del reconocimiento y el contagio de las emociones que produce la conectividad instantánea a la corriente de opinión del grupo al que el sujeto se siente afín. Pues bien, estos tres elementos forman un cóctel explosivo que rompe con el corsé (epistémico y ético) que supone la razón comunicativa para la identidad grupal en la red. Por decirlo de una manera más clara: el uso expresivo de la comunicación en busca de reconocimiento y el contagio emocional propiciado por la red rompen con el deber de civilidad que Rawls (1996) creía necesario en la razón pública. En concreto, estos elementos quiebran el deber de sinceridad por el que los sujetos deben apoyar algo en el debate público porque lo creen realmente y deben dar las razones reales por las que apoyan esa propuesta (Arias Maldonado, 2016). El deber rawlsiano de sinceridad se quiebra porque en la red muchos sujetos actúan motivados por el intento de mantener su estatus dentro del grupo. Esta necesidad de estatus hace que muchos sujetos, guiados por su identidad grupal, expongan contenidos o actúen simplemente con el objetivo de obtener notoriedad, sin importar si los sujetos creen sinceramente lo que están publicando. De hecho, los comentarios más exagerados (y, por ende, insinceros) son los que generan más popularidad. Particularmente, los estudios revelan que los tuits que contienen expresiones

relativas a emociones morales como la indignación o la culpa obtienen un 20% más de retuits que aquellos otros que no apelan a emociones morales (Brady, Crockett y Van Bavel, 2020).

Existen diversos ejemplos de esos usos expresivos de la red en los que los sujetos actúan guiados por su identidad grupal y por la obtención de notoriedad o por un mero desahogo emocional. Los más relevantes son: los discursos incendiarios, los linchamientos virtuales y el exhibicionismo moral. Como trataré de defender a continuación, estos usos de la red no generan una polarización real de posiciones de los usuarios. En ellos se produce una polarización artificial, es decir, una competencia por la notoriedad mediante un uso expresivo de la comunicación que genera una impresión ilusoria de polarización o indignación.

Discursos incendiarios

Un primer ejemplo son los discursos incendiarios dirigidos contra los adversarios ideológicos⁶. Son muchos los tuiteros, *youtubers* o blogueros que publican contenidos con la finalidad de enardecer a sus seguidores y provocar también la reacción airada de sus adversarios. Los discursos incendiarios son, por definición, contenidos hiperbólicos en los que se dibuja una caricatura del oponente. No se le presenta como un adversario legítimo sino como una amenaza existencial.

En todo caso, estos contenidos se publican con un sentido cáustico e irónico, puramente expresivo y por mero interés reputacional. Es, así, un relato exagerado y, por tanto, insincero. Buen ejemplo de ello son los tuiteros políticos, caracterizados por su ingenio para sintetizar en unos pocos caracteres una crítica mordaz al adversario y por hacerlo con un tono hiriente para este, lo que desata la simpatía de sus seguidores. Pero la fama de estos tuiteros proviene justamente de que sus seguidores reconocen el sentido puramente sarcástico de sus tuits. La gente no les sigue para estar informada de la actualidad política, sino para reírse de las caricaturas que esos tuiteros hacen de los adversarios políticos.

Pues bien, este sentido hiperbólico que caracteriza a los discursos incendiarios, que es fácilmente reconocible con los tuiteros, debemos aplicarlo también a otros contenidos publicados por ciertos *youtubers* o blogueros. Muchos de ellos difunden un tipo de discurso incendiario que bien podemos denominar pornografía epistémica. Este concepto, acuñado por Andrew Spear (2020), hace referencia a la información presentada generalmente en forma de meme, tuit, vídeo o artículo que expresa un punto de vista partidista de un

⁶ Hablo de discursos incendiarios y no de discursos de odio por dos motivos: porque los discursos de odio suelen estar dirigidos a minorías tradicionalmente oprimidas y, segundo, porque los discursos de odio, a diferencia de los discursos incendiarios, no tienen un carácter meramente expresivo o retórico, sino que llaman realmente al rechazo de ciertos colectivos a través de su deshumanización (Cortina, 2017).

problema con el objetivo de activar los sesgos cognitivos de los miembros del grupo, como el sesgo de confirmación o el sesgo intragrupo/extragrupo. Es decir, la pornografía epistémica es un tipo de contenido sesgado y parcial que apela al deseo básico de los sujetos de ver validadas sus creencias y sentir que pertenecen a una comunidad que comparte sus mismas creencias.

La pornografía epistémica no se define por el contenido (sesgado) que difunde, sino por el objetivo que persigue: excitar los deseos básicos de la persona, al igual que sucede con la pornografía sexual. Como señala Spear (2020), se trata de un contenido pornográfico por tres motivos: primeramente, porque trata de atraer la atención de los miembros del grupo activando su naturaleza básica. Segundo, porque esa apelación a los deseos básicos atrapa a los sujetos confirmando su cosmovisión y causándoles una sensación de disfrute. Esta sensación impulsa a los sujetos a compartir esos contenidos con otros usuarios, potenciando su viralidad. Y, en tercer lugar, por su carácter ficcional. Así como la pornografía sexual activa la naturaleza sexual de los sujetos, pero nadie cree que viendo pornografía conoce todo lo relativo a las relaciones románticas y sexuales, del mismo modo, el consumo de pornografía epistémica no debería llevar a nadie a pensar que tiene una concepción adecuada de la actualidad política.

Un ejemplo de pornografía epistémica son los libelos dirigidos a desacreditar las propuestas políticas de los adversarios mediante falacias como la del hombre de paja. Son discursos que, en lugar de cuestionar el contenido de esa propuesta, la asocian con ideologías extremas, con regímenes fracasados o con líderes totalitarios del pasado. Estos son discursos compuestos de expresiones inflamantes, epítetos descalificatorios, hipérboles y acusaciones infundadas que conforman una caricatura de esas propuestas y, por ende, de sus proponentes. Con esta caricatura los autores tratan de provocar en la audiencia el rechazo irreflexivo tanto de esas propuestas como de sus defensores. Como resultado, se pretende reforzar la buena imagen del propio grupo por contraste con el estereotipo hecho de los adversarios.

La mera exposición de los sujetos a estos contenidos que caricaturizan al adversario tiene consecuencias a nivel social. Por un lado, aumenta la denominada percepción de polarización, es decir, la tendencia de los sujetos a exagerar las diferencias entre ellos y sus oponentes. Las personas, por el sesgo de disponibilidad, piensan que el retrato hecho del oponente en estos medios es un estereotipo real de los adversarios (Hwang, Kim y Huh, 2014). Pero, sobre todo, si el individuo no capta el carácter puramente retórico de esos discursos, ello puede provocar un aumento de la polarización afectiva: al tener en su mente una imagen desfigurada de sus adversarios, las personas pueden generar mayores sentimientos de rechazo hacia ellos.

Para evitar que esto suceda es fundamental comprender que esos contenidos incendiarios contra los adversarios están escritos por un mero beneficio reputacional, y para excitar los sentimientos tribales de los miembros del grupo (y de los adversarios).

Estos discursos, gracias a su viralización, provocarán el sentimiento de indignación en la audiencia y un reforzamiento de la imagen negativa de los adversarios. Ambos elementos, la sensación de indignación compartida por la audiencia y el estereotipo creado del adversario como una amenaza para el grupo, crearán entre la audiencia la sensación errónea de que hay una gran separación entre ellos y los adversarios, lo que hará parecer que la polarización social es elevada. Pero esta sensación está por completo alejada de la realidad, al basarse en la inflamación emocional provocada por un relato hiperbólico que desfigura la realidad del adversario (Brady, Crockett y Van Bavel, 2020).

Esto es precisamente el fenómeno de la polarización artificial: la creación de una percepción errónea de polarización provocada por un uso de internet que busca meramente mejorar la posición de los usuarios en su grupo a través de la condena visceral del adversario. Se trata de una polarización artificial porque ni quienes han provocado esa atmósfera de hostilidad con la publicación de contenidos denigratorios creen esas afirmaciones (pues las hacen por mero beneficio reputacional), ni, por tanto, el contenido de esos mensajes se ajusta a unos principios básicos de verdad, veracidad y justicia que deben regir el discurso público. Esos contenidos expresivos producen únicamente una falsa sensación de polarización y una imagen deformada del adversario, pero no un reforzamiento de las posiciones de los sujetos, ya que, como señalé en la sección "Formas de polarización", la polarización de posiciones presupone el uso de la razón comunicativa. De ahí que sea necesario denunciar el carácter artificial de esta polarización. Ello permitirá evitar que quienes no reconocen la naturaleza puramente hiperbólica y provocadora de esos contenidos puedan polarizar sus posiciones al visualizarlos.

Linchamientos virtuales

Esta polarización artificial, fruto de la expresión de la identidad grupal mediante un uso expresivo de la red, se produce también en usos activos de internet. Una muestra de ello es la participación en linchamientos virtuales, conocidos también como tormentas de basura (Han, 2014) o tormentas de fuego (Pfeffer, Zorbach y Carley, 2014). Este fenómeno consiste en la descarga repentina de una gran cantidad de mensajes denigratorios contra una persona o grupo en las redes sociales.

Los linchamientos virtuales vienen provocados por la distorsión que el mundo virtual provoca en la comunicación. Cuando las personas con pocos seguidores ponen un comentario en sus redes sociales pueden pensar que lo hacen para su grupo de amigos. Así un comentario jocoso introducido por alguien en mitad de una conversación en Twitter sobre cualquier asunto puede recibir la aprobación de sus seguidores en forma de un retuit. Sin embargo, internet rompe la distinción entre el lenguaje público y el lenguaje privado, ya que retuietar ese comentario como forma de aprobación supone sacarlo de su contexto comunicativo y proyectarlo sobre una nueva audiencia ajena a la conversación inicial. Esta

eliminación del contexto automáticamente transforma lo que era una broma dirigida a unos amigos en una afirmación ofensiva que viola las convicciones más firmes de la sociedad y convierte a su autor en un ser abominable, merecedor de la reprobación pública (Soto Ivars, 2017)⁷.

Para entender la naturaleza de los linchamientos virtuales, algunos autores trazan un paralelo con las antiguas cazas de brujas (Johnen, Jungblut y Ziegele, 2018; Haidt y Lukianoff, 2019). Ambos fenómenos comparten cinco rasgos: El primero es que, al igual que las cazas de brujas, los linchamientos virtuales son referidos a delitos contra el colectivo. La denuncia atrae la atención y desata el pánico entre los demás miembros, haciendo que perciban al acusado como una amenaza para la estabilidad grupal. En segundo lugar, ambos fenómenos se caracterizan por la volatilidad. Las cazas de brujas surgían de repente, causadas por malentendidos o habladurías, se desarrollaban por un periodo de tiempo y luego desaparecían. De un modo análogo, pero más fugaz, las tormentas de fuego surgen de manera repentina, por una acusación aparentemente trivial, y se esfuman de repente, a las pocas horas, cuando el conjunto de la masa ha descargado su ira contra la víctima. El tercer elemento de semejanza entre las cazas de brujas y los linchamientos virtuales es su carácter ritual. El linchamiento constituye toda una ceremonia colectiva por la que el grupo proyecta una violencia moralista contra un agresor de las normas del grupo. Ese carácter ritual se refleja en el hecho de que nadie ose introducir proporcionalidad, discutir la veracidad de la acusación ni intente defender al linchado. Todo el mundo es consciente de que hacer eso supondría interrumpir lo que es en realidad un ritual colectivo. Los linchadores, lejos de atender a sus argumentos, le acusarían de colaborar con el linchado y acabarían linchándolo también. Este carácter ritualista explica los dos últimos rasgos en común entre las cazas de brujas y los linchamientos virtuales. El cuarto es que los linchamientos virtuales, como las cazas de brujas, no son un método de impartir justicia, sino un medio de control social. El instigador del linchamiento aprovecha la denuncia para demostrar ante los demás su especial sensibilidad moral al detectar la violación de una norma del grupo. Pero la denuncia también sirve para poner a prueba la lealtad del resto de los miembros a esos valores grupales, incitando su adhesión a esa denuncia. Esto explica el quinto y último elemento: la desproporcionalidad entre la acusación realizada y las consecuencias del linchamiento. Acusaciones por emplear términos inapropiados en una conversación privada hecha pública pueden acabar con el

⁷ Un conocido ejemplo de linchamiento virtual es el sufrido por Justine Sacco en 2013. Esta joven antes de subir a un avión en EE. UU. con destino a Sudáfrica puso en Twitter un desafortunado comentario con tintes racistas. Durante las doce horas del vuelo ese comentario se viralizó y se convirtió en *trending topic*. Al aterrizar en Ciudad del Cabo, cuando Justine encendió su móvil descubrió que su vida había sido destruida. Los linchadores habían conseguido que la despidieran del trabajo y había periodistas esperándola en el aeropuerto. Aunque era una persona completamente desconocida, un comentario suyo en Twitter dirigido a sus escasos contactos la había convertido a ojos de los usuarios de esa red en un monstruo racista objeto del escarnio mundial (Murray, 2020).

trabajo o la reputación de una persona. Esto se produce porque el fin de los linchadores no es realmente castigar al infractor, sino exhibir su adhesión a los valores del grupo.

Pues bien, estos rasgos de los linchamientos virtuales nos ayudan a entender los motivos por los que las personas se unen a ellos. Estudios realizados reflejan que las personas no se unen a un linchamiento virtual porque tengan un conocimiento del caso en cuestión que los lleve a indignarse y a querer manifestar su indignación en la red. Justamente porque el linchamiento virtual, como la caza de brujas, es una acción colectiva, el elemento clave que lleva a las personas a unirse es la adopción de la identidad colectiva: los sujetos desean sentirse parte del grupo (Gruber, Mayer y Einwiller, 2020). Así, los sujetos se unen al linchamiento para satisfacer un deseo de formar parte de un grupo que hace algo unido. No en vano, es como miembros del colectivo como los sujetos reconocen ese comportamiento denunciado como una amenaza para la estabilidad del grupo⁸.

Por otro lado, la denuncia realizada activa en el resto de los miembros el sentimiento de comunidad al ver amenazada la estabilidad grupal. Es esa amenaza, y no el conocimiento real del caso, lo que despierta en el grupo la emoción moral de indignación contra la víctima del linchamiento. Esa indignación colectiva une al grupo en la condena de esa persona mediante una catarata de comentarios expresivos de carácter denigratorio y que da como resultado un aparente clima de polarización intragrupal⁹. Sin embargo, como señala Han (2014), esa sensación de indignación es consecuencia del contagio emocional producido por la viralización de la denuncia por parte de los sujetos para satisfacer su deseo de pertenencia. Es decir, los sujetos se suman al linchamiento virtual con una finalidad expresiva: dar cauce a un sentimiento de indignación provocado por el impacto sufrido por la visualización del mensaje recibido, no porque estén realmente indignados con ese hecho denunciado (ya que no lo conocen en detalle). Siguiendo el marco apuntado desde la neuroética, la participación en el linchamiento virtual constituye una mera *reacción expresiva* emocional de un sentimiento de indignación contagiado por la exposición al contenido recibido, no una *acción decidida* a partir de un juicio moral de indignación (Monin, Pizarro y Beer, 2007). Para expresarlo en términos aristotélicos, la acción decidida requiere el uso de la razón. En cambio, los linchamientos virtuales consisten en una descarga de reacciones afectivas provocada por la visualización de un contenido impactante.

⁸ En cambio, visto desde fuera del grupo, ese acto que recibe la condena del grupo resulta una acción sin importancia. Por eso también la respuesta del grupo hacia quien introduzca proporcionalidad desde fuera es violenta, porque esa crítica se percibe como un ataque a los valores sagrados para la identidad grupal.

⁹ Los linchamientos virtuales se confunden con la polarización real debido a la falsa percepción de consenso que generan: quienes adoptan la identidad grupal perciben un ataque al grupo, les embarga un sentimiento de indignación y participan del linchamiento. En cambio, quienes permanecen ajenos a ese ritual colectivo quedan callados para evitar ser linchados también. En consecuencia, desaparece cualquier voz discrepante, creándose un falso consenso en el grupo y un aparente reforzamiento de posiciones.

Otros autores citan también el deseo de reconocimiento social como motivo que lleva a las personas a unirse a un linchamiento virtual. La persona desea mejorar su estatus dentro del grupo, y lo hace mostrando públicamente su adhesión al grupo sumándose al linchamiento. Los *likes* o retuits que reciba su comentario por parte de otros usuarios aumentarán su respetabilidad en el grupo (Johnen, Jungblut y Ziegele, 2018). Pero, de este modo, lo que lleva a los individuos a sumarse al linchamiento no es una indignación real fruto del conocimiento del caso, sino el deseo de mejorar la posición dentro del grupo, así como la respuesta emocional al impacto sufrido por la visualización del mensaje recibido. Por eso, mantengo que el linchamiento virtual no representa una muestra de polarización real de posiciones de los sujetos, sino de polarización artificial: un agregado de mensajes expresivos de rechazo hacia la víctima del linchamiento realizados por un mero interés reputacional, que genera una falsa sensación de indignación entre la audiencia.

Exhibicionismo moral

Un último uso de internet que provoca una polarización artificial es el exhibicionismo moral – esto es, la expresión de indignación moral por algún asunto que alguien realiza con el único objetivo de que los demás reconozcan su especial sensibilidad moral (Tosi y Warmke, 2016). Cuando este fenómeno se produce entre individuos de grupos rivales, el exhibicionista moral busca presentarse como alguien moralmente superior al adversario, al exhibir una sensibilidad moral más fina. Con ello pretende conseguir que sus demandas morales tengan una hegemonía en el discurso público frente a las demandas del adversario. Pero cuando el exhibicionismo tiene lugar entre personas afines se desata una dinámica similar a la polarización grupal descrita por Sunstein, pero que, en realidad, como paso a mostrar, constituye una polarización artificial. En concreto, tiene lugar una escalada condenatoria en la cual, si alguien denuncia que a una persona hay que censurarla públicamente por violar una norma del grupo, otro dirá que deberían despedirla, otro dirá que hay que denunciarla judicialmente, etc.

Ahora bien, esta escalada condenatoria no se produce como fruto de un proceso de polarización de posiciones dentro del grupo. Más bien es consecuencia de la naturaleza puramente expresiva del exhibicionismo moral. Dado que el grado de indignación se considera indicativo de contar con un sentido moral más desarrollado, aquel que exprese una mayor indignación ocupará una mejor posición dentro del grupo. Por ese motivo se inicia la escalada condenatoria en la que todos compiten por parecer más indignados y ganar con ello el respeto de todos. Esta dinámica se explica también porque en muchos contextos actuales la valoración moral por parte del grupo se ha desplazado de la intención real que tenía el autor del acto a la indignación expresada por la víctima o por el denunciante de la acción (Haidt y Lukianoff, 2019).

Así, en el exhibicionismo moral nos encontramos puramente ante una competición por el estatus. Efectivamente, la motivación que guía a los exhibicionistas morales no es sino mostrar una mayor sensibilidad ética que sus afines para, con ello, ascender de posición en el grupo. Esta motivación rompe las pretensiones de validez de la comunicación relativas a la verdad, la veracidad y la justicia, ya que el hablante no busca decir verdad (aunque realmente exprese una indignación sincera por un hecho) sino impresionar a la audiencia. Conscientes de ello, el resto de los sujetos inician una alocada carrera de expresiones gratuitas de indignación realizadas con el único fin de mejorar su posición en el grupo y expresar un desahogo emocional ante el impacto del mensaje recibido. Esto inevitablemente lleva a los sujetos a decir cosas que no creen. Porque lo que buscan es quedar por encima del resto en el pódium de la indignación moral. Esto explica que se produzca una escalada condenatoria de afirmaciones insinceras que da como resultado una burbuja artificial de indignación inflamada por la dinámica emotiva y vilipendiada de las redes (Grubbs et al., 2019). En este sentido, el exhibicionismo moral entre afines, lejos de propiciar una polarización de posiciones dentro del grupo, representa lo que llamo polarización artificial: una sucesión de afirmaciones arbitrarias de indignación sobre un asunto realizadas con el objetivo de mejorar la posición de los usuarios en el grupo y que da como resultado en la audiencia una sensación ficticia de que el grupo ha polarizado (es decir, radicalizado o reforzado) su posición condenatoria sobre ese asunto.

Evidentemente, a esto se podría decir que esta polarización artificial tiene consecuencias reales sobre las personas, especialmente sobre las víctimas de los linchamientos virtuales, que pueden ver arruinada su reputación o perder su empleo. Pero justamente por eso se hace más necesario, en primer lugar, discriminar entre los procesos de polarización real y la polarización artificial. Es decir, hay que reconocer los supuestos dialógicos de ciertos usos de internet que provocan una polarización, ya sea intragrupal o intergrupal, de posturas, y, por tanto, una mayor división de la sociedad, como he mostrado en la sección "Formas de polarización". Ello permitirá distinguir estos usos de la red de otros usos de internet como los expuestos en esta sección (discursos incendiarios, linchamientos virtuales y exhibicionismo moral), en los que se produce una polarización artificial. De este modo será posible advertir que estos fenómenos del linchamiento virtual, los discursos incendiarios o el exhibicionismo moral no provocan una polarización real de la sociedad, pues consisten en una sucesión de discursos expresivos que emplean los usuarios en la red para obtener notoriedad explotando la indignación en su audiencia. Estos usos de internet producen una sensación ilusoria de confrontación en la sociedad. Es una sensación ilusoria porque es fruto de expresiones insinceras de indignación y porque presenta un retrato alejado de cómo son realmente los adversarios. Así, para evitar las consecuencias perversas que puede arrojar la polarización artificial a nivel social, en la última sección ofreceré unas breves indicaciones para terciar con este fenómeno.

Cómo frenar la polarización artificial

Evitar la polarización, real y artificial, en el debate público debe ser un objetivo en toda reflexión sobre la polarización. En un artículo anterior (Pérez Zafrilla, 2017) ya abordé las estrategias para afrontar los procesos reales de polarización, surgidos en el debate público. Estas iban en la línea de reivindicar las virtudes cívicas desarrolladas desde la democracia deliberativa (Gutmann y Thompson, 1997), como: el respeto mutuo (no dudar de la actitud ética del interlocutor); la integridad cívica (defender lo que sinceramente se cree); la caridad intelectual (presumir la racionalidad y sinceridad del otro); la magnanimidad cívica (reconocer el estatus moral de las posiciones opuestas); tener una mente abierta para revisar nuestras creencias a la luz de los argumentos presentados por otros; y, por último, economizar el desacuerdo, esto es, buscar los puntos de convergencia con el otro y evitar el desacuerdo innecesario.

Sin embargo, combatir la polarización artificial, dado que responde a un uso expresivo de la red en el que no se produce un diálogo sincero entre las personas, requiere en mi opinión, del recurso a dos virtudes diferentes que ya fueran apuntadas por Adela Cortina para afrontar la dinámica de la sociedad consumista (Cortina, 2003) y la convivencia política (Cortina, 2013): la lucidez y la cordura. Estas virtudes permitirían también navegar en la red sin caer en formas de polarización artificial.

La lucidez permite a la persona, en primer lugar, reconocer los distintos tipos de contenido que hay en la red. Igual que en la televisión se distinguen los formatos de la información (informativos) y del entretenimiento (concursos), así también el usuario de la red debería ser capaz de discriminar cuándo una página de web o canal *online* ofrece un contenido fiable y cuándo ofrece mera pornografía epistémica. Esta última puede serle útil para entretenerse, pero no debería emplearla para informarse sobre la actualidad.

Segundo, la lucidez permite reconocer la tipología de perfiles que podemos encontrarnos en la red. Así como hay usuarios reales que buscan opinar sobre un tema, también hay otros, los trols, que pretenden ridiculizar a su interlocutor ante su audiencia para conseguir la aprobación de sus seguidores. También hay perfiles tras los cuales no hay personas reales sino bots que replican mensajes automáticamente. Conocer esta tipología permitirá advertir en qué contextos la interacción con otros usuarios en la red da lugar a verdaderos procesos de comunicación y cuándo es mejor no perder el tiempo interactuando con determinados perfiles.

Tercero, la lucidez ayuda a reconocer el ecosistema de la red. Por ejemplo, permite saber que ciertos temas son tendencia en la red no por su relevancia real para la ciudadanía, sino porque están siendo amplificados por cuentas falsas y bots al servicio de grupos de presión o simplemente porque es un contenido patrocinado por una empresa. Asimismo, permite tomar conciencia de que los contenidos incendiarios publicados en ciertos medios tienen una función meramente expresiva y retórica, pretenden explotar

nuestras emociones más básicas mediante el uso de expresiones inflamantes y están dirigidos a mejorar la reputación de su autor dentro de su tribu política.

Por último, a nivel individual, la lucidez ayuda a los sujetos a desentrañar los motivos por los que consumen discursos incendiarios o participan en linchamientos virtuales o el exhibicionismo moral. Lo hacen debido a los sesgos cognitivos de confirmación o de intragrupo/extragrupo activados cuando las personas adoptan su identidad grupal. También lo hacen para mostrar en el escaparate de la red su adhesión a la identidad grupal o para dar expresión a un sentimiento de indignación contagiado por la visualización de ciertos contenidos.

Por su parte, la cordura permite actuar en la red en consecuencia a lo descubierto con la virtud de la lucidez. De manera general, la cordura, como trasunto cordial de la virtud clásica de la prudencia, permite a los ciudadanos actuar con la debida proporcionalidad y mesura en los diferentes contextos digitales:

Por un lado, permite tomar con la debida ironía los discursos incendiarios, igual que hacen los ciudadanos lúcidos con los tuits de los tuiteros satíricos, las tertulias políticas o las homilías partidistas de los comunicadores radiofónicos. Si esos contenidos son meras construcciones retóricas plagadas de soflamas y expresiones inflamantes que pretenden excitar las emociones tribales de odio al adversario, el ciudadano lúcido debe reaccionar ante esos discursos con el correspondiente cinismo. Si el creador de esos discursos ni siquiera cree en ellos, sino que los construye por mero afán de prestigio, ¿qué sentido tiene que los usuarios tomen en serio esas representaciones hiperbólicas pensando que les revelan una verdad sobre el adversario ideológico? Además, esos discursos presuponen un modelo de ciudadanía entendida como masa: conciben a los ciudadanos como sujetos anómicos y tribalistas, sin un criterio propio y racional, que se mueven por meras reacciones emocionales provocadas por el contenido inflamante de esos discursos. Es decir, los discursos incendiarios tratan a los ciudadanos como menores de edad a los que se puede manipular emocionalmente. En consecuencia, quienes toman en serio esos discursos abdican de su condición de ciudadanos y se convierten en siervos a merced de los poderosos que les marcan lo que deben creer a través de la propaganda transmitida en esos discursos.

En segundo lugar, y en relación con lo anterior, la cordura debe llevar a rechazar el mismo hecho de que en un grupo se pueda ganar reputación mediante la denigración, siquiera retórica, de otras personas o grupos, como el adversario ideológico. Este es un punto fundamental. La convivencia democrática sólo puede construirse desde el reconocimiento del otro como ser digno de respeto, así como desde la forja de un vínculo entre los ciudadanos que lleve a alcanzar objetivos comunes, respetando las diferencias legítimas (Cortina, 2013). En consecuencia, al igual que sucede en las relaciones *offline*, el adversario ideológico debe ser reconocido en la red como alguien con una posición moral legítima y con quien es posible descubrir qué es lo justo a través del diálogo, no como la

representación del mal. Por ese motivo, los discursos incendiarios, a pesar de su carácter ficcional, comparten con los discursos de odio el socavar la convivencia democrática, ya que presuponen una política basada en la confrontación con el enemigo en lugar de en el diálogo con el diferente para construir un futuro compartido (Cortina, 2017).

Así pues, dado que la democracia se asienta sobre el reconocimiento del otro como ser digno y sobre el diálogo entre diferentes para forjar un futuro común, la cordura impulsa a desterrar cualquier forma de discurso que deshumanice al adversario o vaya dirigido a la construcción de trincheras que quiebren la amistad cívica entre las personas. La cordura nos ayuda a reconocer que los discursos que dividen y enfrentan deben ser objeto de reproche en lugar de recibir el elogio a través de *likes*, *retuits* o seguidores, justamente por el daño que hacen a la convivencia. Por eso mismo también, en una sociedad lúcida, que basa la política en la forja de una voluntad común en lugar de en el enfrentamiento, no tendrá cabida tampoco buscar la reputación en la red mediante la denigración de otras personas o grupos, como el adversario.

Tercero, la cordura ayuda a introducir la proporcionalidad al juzgar los hechos y evaluarlos en la medida en que se tiene conocimiento de aspectos elementales del caso, como las intenciones reales de la persona linchada en los casos del linchamiento virtual. También la cordura predispone, en la línea de lo propuesto por Haidt y Lukianoff (2019), a interpretar los actos y palabras de los demás de la forma más caritativa. Esto llevará a los ciudadanos lúcidos a abstenerse de sumarse a linchamientos de personas, pues acarrearán consecuencias desproporcionadas sobre sus víctimas.

Finalmente, la cordura desactiva por completo los actos de exhibicionismo moral. Desde el mismo momento en que los sujetos reconocen el carácter meramente expresivo, insincero y puramente narcisista del exhibicionismo moral, la única respuesta coherente que cabe dar es la negativa a participar en esa escalada, así como el rechazo a quienes comulgan con esa práctica.

Una sociedad de ciudadanos lúcidos permitirá, así, frenar los procesos de polarización artificial en la red y crear un ambiente más calmado al saber distinguir los distintos usos que los sujetos hacen de la red y los motivos que les llevan a hacerlos.

Conclusión

Como he mostrado a lo largo del artículo, la polarización política es un fenómeno que tiene lugar tanto *online* como *offline*. He expuesto las formas en que este proceso se desarrolla, así como las bases evolutivas y los supuestos epistémicos y morales referidos a la comunicación que hacen posible la polarización. Todo ello ha permitido fijar el marco desde el que proyectar el concepto de polarización artificial. Esta idea hace referencia a aquellos usos de internet puramente expresivos en los que la competición por obtener

notoriedad mediante la explotación de la indignación en la audiencia da como resultado una impresión ficticia de polarización.

Comprender la diferencia entre la polarización real de posiciones y la polarización artificial resulta clave para analizar adecuadamente el fenómeno de la polarización en la red, así como los usos que los sujetos hacen de internet. Del mismo modo, la idea de polarización artificial contribuye a forjar ciudadanos lúcidos, capaces de discriminar los diferentes contenidos de la red, así como de desentrañar los motivos que les llevan a consumir ciertos contenidos ideológicos o a participar en acciones como linchamientos virtuales o el exhibicionismo moral. Todo ello permitirá que una ciudadanía lúcida pueda navegar por la red con cordura, evitando conceder una credibilidad indebida a contenidos incendiarios o participar en acciones que puedan dañar a otras personas. Será posible así crear un ambiente en la red más sereno, que desinflame el debate público y encauce el acuerdo político entre diferentes.

Referencias bibliográficas

ABRAMOWITZ, A.; SAUNDERS, K. E. "Is polarization a myth?". *The Journal of Politics*, vol. 70, nº 2, p. 542-555, 2008.

AGRAWAL, A. "What do social media algorithms mean for you?". *Forbes*, 20 abr. 2016. Disponible en: <<https://www.forbes.com/sites/ajagrawal/2016/04/20/what-do-social-media-algorithms-mean-for-you/>>. Acceso en: 13 jun. 2021.

ALEXANDER, R. *The biology of moral systems*. New York: Routledge, 1987.

ALTER, A. *Irresistible*. ¿Quién nos ha convertido en yonquis tecnológicos? Barcelona: Paidós, 2018.

ARIAS MALDONADO, M. *La democracia sentimental*. Política y emociones en el siglo XXI. Barcelona: Página Indómita, 2016.

BASIL, C. A., et al. "Exposure to opposing views on social media can increase political polarization". *PNAS*, vol. 15, p. 9.216-9.221, 2018.

BOXELL, L.; GEINTZKOW, M.; SHAPIRO, J. M. "Greater Internet use is not associated with faster growth in political polarization among US Demographic Groups". *PNAS*, vol. 114, nº 40, p. 10.612-17, 2017.

BRADY, W.; CROCKETT, M. J.; VAN BAVEL, J. "The MAD model of moral contagion: the role of motivation, attention, and design in the spread of moralized content online". *Perspectives on Psychological Science*, vol. 15, p. 978-1.010, 2020.

BRENNAN, J. *Contra la democracia*. Madrid: Deusto, 2018.

CHO, J., et al. "Do search algorithms endanger democracy? An experimental investigation of algorithm effects on political polarization". *Journal of Broadcasting & Electronic Media*, vol. 64, nº 2, p. 150-72, 2020.

CORTINA, A. *Por una ética del consumo*. Madrid: Taurus, 2003.

_____. *Ética de la razón cordial*. Oviedo: Nobel, 2007.

_____. *¿Para qué sirve realmente la ética?* Barcelona: Paidós, 2013.

_____. *Aporofobia, el rechazo al pobre.* Barcelona: Paidós, 2017.

FLAXMAN, S.; GOEL, S.; RAO J. M. "Filter bubbles, echo chambers and online news consumption". *Public Opinion Quarterly*, vol. 80, p. 298-320, 2016.

GREENE, J. *Moral Tribes: emotion, reason, and the gap between us and them.* London: Atlantic Books, 2013.

GRUBBS, J. B., et al. "Moral grandstanding in public discourse: status-seeking motives as a potential explanatory mechanism in predicting conflict". *PLoS ONE*, vol. 14, nº 10, e0223749, 2019.

GRUBER, M.; MAYER, C.; EINWILLER, S. "What drives people to participate in online firestorms?". *Online Information Review*, vol. 44, nº 3, p. 563-81, 2020.

GUTMANN, A; THOMPSON, D. *Democracy and disagreement.* Cambridge: Belknap Press of Harvard University Press, 1997.

HABERMAS, J. *Teoría de la acción comunicativa.* Madrid: Taurus, 1987.

_____. *Facticidad y validez.* Madrid: Trotta, 1998.

HAIDT, J. *La mente de los justos: por qué la política y la religión dividen a la gente sensata.* Madrid: Deusto, 2019.

HAIDT, J.; LUKIANOFF, G. *La transformación de la mente moderna.* Madrid: Deusto, 2019.

HAIDT, J.; ROSE-STOCKWELL, T. "The dark psychology of social networks. Why it feels like everything is going haywire". *The Atlantic*, diciembre 2019. Disponible en: <<https://www.theatlantic.com/magazine/archive/2019/12/social-media-democracy/600763/>>. Acceso en: 12 sept. 2020.

HAN, B. C. *En el enjambre.* Barcelona: Herder, 2014.

HILL, R. K. "What an algorithm is". *Philosophy & Technology*, vol. 29, nº 1, p. 35-59, 2015.

HWANG, H.; KIM, Y.; HUH, C. "Seeing is believing: effects of uncivil online debate on political polarization and expectations of deliberation". *Journal of Broadcasting & Electronic Media*, vol. 58, nº 4, p. 621-33, 2014.

IPSOS MORI SOCIAL RESEARCH INSTITUTE. *BBC Global Survey. A World Divided?*, 2018. Disponible en: <<https://www.ipsos.com/sites/default/files/ct/news/documents/2018-04/bbc-global-survey-a-world-divided-2018.pdf>>. Acceso en: 14 jul. 2020.

IYENGAR, S; WESTWOOD, S. "Fear and loathing across party lines: new evidence on group polarization". *American Journal of Political Science*, vol. 59, nº 3, p. 690-707, 2015.

JOHNEN, M.; JUNGBLUT, M.; ZIEGELE, M. "The digital outcry: what incites participation behavior in an online firestorm?". *New Media & Society*, vol. 20, nº 9, p. 3140-3160, 2018.

KIM, Y. "How do news frames influence mass political polarization?". Tesis de Doctorado en Filosofía. University of Alabama, Tuscaloosa, 2016.

KUNDA, Z. "The case for motivated reasoning". *Psychological Bulletin*, vol. 108, nº 3, p. 480-98, 1990.

MCCOY, J.; RAHMAN, T.; SOMER, M. "Polarization and the global crisis of democracy: common patterns, dynamics, and pernicious consequences for democratic policies". *American Behavioral Scientist*, vol. 62, nº 1, p. 16-42, 2018.

MERCIER, H.; SPERBER, D. "Why do humans reason? Arguments for an argumentative theory". *Behavioral and Brain Sciences*, vol. 34, nº 2, p. 57-111, 2011.

MÖLLER, J., et al. "Do not blame on the algorithm: an empirical assessment of multiple recommender systems and their impact on content diversity". *Information, Communication & Society*, vol. 21, nº 7, p. 959-77, 2018.

MONIN, B.; PIZARRO, D. A.; BEER, J. S. "Deciding versus reacting: conceptions of moral judgments and the reason-affect debate". *Review of General Psychology*, vol. 11, nº 2, p. 99-111, 2007.

MURRAY, D. *La masa enfurecida: cómo las políticas de identidad llevaron al mundo a la locura*. Barcelona: Península, 2020.

MUTZ, D. *Hearing the other side: deliberative versus participatory democracy*. Cambridge: Cambridge University Press, 2006.

O'NEIL, C. *Armas de destrucción matemática: cómo el big data aumenta la desigualdad y amenaza la democracia*. Madrid: Capitán Swing Libros, 2017.

PARISER, E. *El filtro burbuja: cómo la red decide lo que leemos y lo que pensamos*. Barcelona: Taurus, 2017.

PÉREZ ZAFRILLA, P. J. "Por qué fracasa la deliberación y cómo podemos remediarlo. Una alternativa ética al enfoque neurocientífico". *Daimon - Revista Internacional de Filosofía*, nº 70, p. 131-146, 2017.

_____. Polarización política: estado de la cuestión y orientaciones para el análisis. In: SANTIBÁÑEZ, C. (ed.). *Emociones, argumentación y argumentos*. Lima: Palestra, 2020.

_____. "Polarización artificial: cómo los discursos expresivos inflaman la percepción de polarización política en internet". *Recerca - Revista de Pensament i Anàlisi*, vol. 26, nº 2, p. 1-23, 2021.

PEW RESEARCH CENTER. "The partisan divide on political values grows even wider". Oct. 2017. Disponible en: <<https://www.pewresearch.org/politics/2017/10/05/the-partisan-divide-on-political-values-grows-even-wider/>>. Acceso en: 12 sept. 2020.

PFEFFER, J.; ZORBACH, T.; CARLEY, M. "Understanding online firestorms: negative word-of-mouth dynamics in social media networks". *Journal of Marketing Communication*, vol. 20, nº 1-2, p. 117-28, 2014.

RAWLS, J. *El liberalismo político*. Barcelona: Crítica, 1996.

RUBIO NÚÑEZ, R. "Los efectos de la posverdad en la democracia". *Revista de Derecho Político*, nº 103, p. 191-228, 2018.

SCHULZ, K. *Being Wrong: adventures in the margin of error*. New York: Portobello Books, 2010.

SHAW, V. F. "The cognitive processes in informal reasoning". *Thinking and Reasoning*, vol. 2, nº 1, p. 51-80, 1996.

SOTO IVARS, J. *Arden las redes: la poscensura y el nuevo mundo virtual*. Madrid: Debate, 2017.

SPEAR, A. "Breaking the epistemic pornography habit: cognitive biases, digital discourse environments and moral exemplars". *Journal of Information, Communication, and Ethics in Society*, vol. 18, nº 1, p. 87-104, 2020.

STROUD, N. J. *Niche News: the politics of news sources*. Oxford: Oxford University Press, 2011.

SUNSTEIN, C. "The law of group polarization". *The Journal of Political Philosophy*, vol. 10, nº 2, p. 175-95, 2002.

_____. *República.com: internet, democracia y libertad*. Barcelona: Paidós, 2003.

TOSI, J.; WARMKE, B. "Moral grandstanding". *Philosophy and Public Affairs*, vol. 44, nº 3, p. 197-217, 2016.

TRIVERS, R. "The evolution of reciprocal altruism". *The Quarterly Review of Biology*, vol. 46, nº 1, 1971, p. 35-57.

WILSON, E. O. *Sociobiología*. Madrid: Omega, 1995.

Abstract

Digital tribalism, between fury and farce: let's burst the bubble of artificial polarization on the Internet

In recent years, public opinion has grown concerned about political polarization. Some authors point to the Internet and social networks as a cause of this polarization. In this context, this article addresses the different uses that subjects make of the Internet and that give rise to different forms of polarization. The article starts from the bibliographic review on the neuropsychological bases of political behavior and the nature of political polarization. On this basis, the article presents the ways in which polarization takes place on the Internet, as well as the moral and epistemic assumptions of political polarization. Faced with these forms of polarization, the article develops the concept of *artificial polarization*. This concept aims to explain how the expressive uses of the network generate a wrong perception of polarization among users. Examples of artificial polarization are flaming, firestorms, and moral grandstanding. As a result, the article presents some pointers to disable the processes of artificial polarization and build a calmer environment on the web.

Keywords: political polarization; artificial polarization; flaming; firestorms; moral grandstanding

Resumo

Tribalismo digital, entre a fúria e a farsa: vamos estourar a bolha da polarização artificial na internet

Nos últimos anos, a opinião pública tem se preocupado com a polarização política. Alguns autores apontam a internet e as redes sociais como causa dessa polarização. Nesse contexto, este artigo aborda os diferentes usos que os indivíduos fazem da internet e que dão origem a diferentes formas de polarização. O artigo parte da revisão bibliográfica sobre as bases neuropsicológicas do comportamento político e a natureza da polarização política. Com base nisso, o artigo apresenta os modos como ocorre a polarização na internet, bem como os pressupostos morais e epistêmicos da polarização política. Diante dessas formas de polarização, o artigo aborda o conceito de *polarização artificial*. Este conceito visa explicar como os usos expressivos da rede geram uma percepção equivocada de polarização entre os usuários. Exemplos de polarização artificial são discursos incendiários, linchamentos virtuais ou postura moral. Como resultado, o artigo apresenta algumas indicações para desativar os processos de polarização artificial e construir um ambiente mais calmo na rede.

Palavras-chave: polarização política; polarização artificial; discursos incendiários; linchamentos virtuais; arrogância moral

Résumé

Tribalismo digital, entre fureur et farce: perçons la bulle de la polarisation artificielle sur internet

Ces dernières années, l'opinion publique s'est inquiétée de la polarisation politique. Certains auteurs désignent Internet et les réseaux sociaux comme une cause de cette polarisation. Dans ce contexte, cet article aborde les différents usages que font les utilisateurs d'Internet et qui donnent lieu à différentes formes de polarisation. L'article part de la revue bibliographique sur les bases neuropsychologiques du comportement politique et la nature de la polarisation politique. Sur cette base, l'article présente les modes de polarisation sur Internet, ainsi que les hypothèses morales et épistémiques de la polarisation politique. Face à ces formes de polarisation, l'article développe le concept de *polarisation artificielle*. Ce concept vise à expliquer comment les usages expressifs du réseau génèrent une fausse perception de la polarisation chez les utilisateurs. Des exemples de polarisation artificielle sont les discours incendiaires, les lynchages virtuels ou les postures morales. En conséquence, l'article présente quelques conseils pour désactiver les processus de polarisation artificielle et créer un environnement plus calme sur le Web.

Mots-clés: polarisation politique; polarisation artificielle; discours incendiaires; lynchages virtuels; arrogance morale

Artigo submetido à publicação em 17 de setembro de 2020.

Versão final aprovada em 9 de novembro de 2021.

Opinião Pública adota a licença Creative Commons CC-BY.

